DOCUMENT RESUME

ED 174 112 HE 011 454

AUTHOR Campbell, S. Duke; Greenberg, Barry
TITLE The Use of Multiple Regression and Trend Analysis to Understand Enrollment Fluctuations. AIR Forum 1979 Paper.
PUB DATE May 79
NOTE 13p.; Paper presented at the Annual Forum of the Association for Institutional Research (19th, San Diego, California, May 13-17, 1979); Appended data (3 pages) may not reproduce well

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Planning; Economic Factors; *Enrollment Projections; Higher Education; Institutional Research; *Mathematical Models; *Multiple Regression Analysis; *Predictor Variables; *Trend Analysis; Unemployment
IDENTIFIERS *AIR Forum 1979; *Florida International University

ABSTRACT

The development of a predictive equation capable of explaining a significant percentage of enrollment variability at Florida International University is described. A model utilizing trend analysis and a multiple regression approach to enrollment forecasting was adapted to investigate enrollment dynamics at the university. Four independent variables were used in a stepwise linear regression analysis to explain variance in the total university full-time-equivalent enrollment. Three of the four variables used are presently employed in business forecasting models to account for changes caused by trends, seasonal variations, and cyclical variations. The fourth variable was introduced to account for the opening and operation of a new campus. The model accounts for 99.6 percent of the variation observed in total university enrollments for the past 20 quarters and provides estimates of the enrollment within one percent of the enrollments actually observed. The model also has the potential for making reasonably accurate and reliable short-range enrollment forecasts. The model demonstrated that changes in economic conditions in the county, reflected by changing unemployment rates, correlate significantly and positively with changes in total university enrollment. Total university enrollment did not change from its second to its sixth year of operation except for those changes that could be accounted for by changes in unemployment in the county. The opening and operation of the new campus increased total university enrollment by about six percent. (SC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED174112

THE ASSOCIATION FOR INSTITUTIONAL RESEARCH

This paper was presented at the Nineteenth Annual Forum of the Association for Institutional Research held at the Town and Country Inn in San Diego, California, May 13-17, 1979. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC collection of Forum papers.

Mary Corcoran
University of Minnesota
(Editor, AIR Forum Publications)

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

AIR

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE USE OF MULTIPLE REGRESSION AND TREND ANALYSIS TO UNDERSTAND ENROLLMENT FLUCTUATIONS

by

S. Duke Campbell

and

Barry Greenberg

Florida International University

Paper presented at the Annual Forum of the
Association for Institutional Research,
San Diego, California, May, 1979.

I. Introduction

The challenge to explain periodic enrollment fluctuations is faced by most institutional researchers in an effort to provide understanding and predictive capability to decision making administrators. This task has, of course, increased in importance in recent years as legislatures have imposed new financial constraints, and moved toward enrollment-driven funding models. At Florida International University, the opening of a new campus provided additional incentive to explore the issue of enrollment planning since it was recognized that the new site would affect student flow throughout the University. This investigation was undertaken to better understand historical enrollment patterns and to develop a predictive equation capable of explaining a significant percentage of enrollment variability so that the challenges described above could be addressed.

II. Methodology

In a paper presented at the 1978 AIR Forum, Salley (1978) utilized trend analysis and a multiple regression approach to enrollment forecasting. His findings and suggestions led the authors to believe that his model, in an adapted format, might assist them in understanding enrollment dynamics at their own institution.

In this investigation, four independent variables were used in a stepwise linear regression analysis to explain variance in the dependent variable (total University FTE enrollment). Three of the four independent variables used are presently employed in business forecasting models to account for changes caused by trends, seasonal variations and cyclical variations (as per Salley's suggestions). The fourth variable was introduced to account for the opening and operation of the North Miami Campus for the last five quarters. In all, data were used for twenty quarters of operation, Fall 1973 to Summer 1978.

The trend variable accounts for changes in enrollment which might result from such things as the addition of new programs, increased publicity, better draw because of an increasingly good reputation and the like. Changes such as these would produce a positive trend (increase in enrollment over time). Similarly of course, negative changes in these and other factors could produce a negative trend (decrease in enrollment over time). The input variable is simply the numerical sequence for the quarters used in the analysis, i.e., Fall 1973 = Quarter 1, Winter 1974 = Quarter 2 . . . Summer 1978 = Quarter 20.

SUMMARY

A linear multiple regression model has been developed which accounts for 99.6% of the variation observed in total University enrollments for the past twenty quarters and provides estimates of the enrollment within 1% of the enrollments actually observed. The model also has the potential for making reasonably accurate and reliable short-range enrollment forecasts.

Specific conclusions drawn from the model relating to Florida International University's specific situation include:

1. Changes in economic conditions within Dade County, as reflected by changing unemployment rates, correlate significantly ($r = .90$) and positively (95 FTE/1% change in unemployment) with changes in total University enrollment.
2. Total University enrollment did not change from its second to its sixth year of operation except for those changes which could be accounted for by changes in Dade County unemployment.
3. The opening and operation of a new campus at North Miami has increased total University enrollment by approximately 368 FTE students (5-6%).

The seasonal variable attempts to account for regular changes which occur within a given, repetitive time frame. Historically, for example, a typical Summer quarter enrollment is less than Fall. Fall enrollment is typically less than Winter, etc. By computing a seasonal index for each quarter, it is possible to seasonally adjust the raw data or deseasonalize the data. The seasonal indices (SI) may be held as a constant for a given quarter or may be adjusted if a trend in the seasonal index is observed. In this analysis a constant seasonal index was used for each quarter. These were: Summer SI = 0.67; Fall SI = 1.14; Winter SI = 1.161; and Spring SI - 1.029. (Note: the quarterly seasonal index is not the same numerically as the quarter/four quarter average ratio; however, because of the way each is computed, they are similar in magnitude.)

The cyclical variable is used in business forecasting to account for a factor which is known to influence (or at least correlates with) the dependent variable in a periodic or cyclical manner. A relationship between University enrollment and the local economy has been found in other settings and suggested for Miami as well. To test this possible relationship, seasonally adjusted enrollments (that is, enrollment with the effect of quarterly variations removed) were regressed against seasonally adjusted employment data for Dade County (Miami SMSA). A significant and positive correlation was found (r^2 = .64). However, the highest correlation (r^2 = .81) was found when unemployment was lead one quarter with respect to enrollment, that is, the prior quarter unemployment correlated better with current quarter enrollment than did current quarter unemployment. Intuitively, this seems reasonable and to be expected.

The North Miami Campus opening appears to have significantly affected University enrollment as can be observed in the seasonally adjusted data coincidental with the quarter the North Miami Campus began regular, full-scale operations (Summer 1977). To account for this change a dummy variable was introduced into the model. The variable has a value of zero for the first fifteen (15) quarters (before North Miami Campus operations) and a value of one (1) for the five quarters of the North Miami operation. This technique permits us to assess both the magnitude and significance of North Miami Campus operations, if any, and of course assumes no other change in University operation occurred coincidentally.

III. Results

Using the total University enrollment data for the past twenty quarters and the corresponding independent

variables described earlier, the multiple linear regression model produced a "best fit" equation for the data observed. The equation was:

$$y = -841.9 - 6.6\ (X_1) + 6357\ (X_2) + 95\ (X_3) + 368.2\ (X_4)$$

Where y = total University enrollment

X_1 = trend factor (quarters expressed numerically)

X_2 = seasonal index for the quarter

X_3 = unemployment rate for Dade County (seasonally adjusted and lead one quarter)

X_4 = dummy variable for the North Miami Campus opening and operation (0 or 1)

The multiple correlation coefficient for the data was 0.998, and the estimated y values were, on the average, within 1% of the observed y values. In short, the model estimates enrollment and accounts for quarterly changes in our enrollment very accurately over the past twenty quarters (see figure 1).

The seasonal variable (regular quarterly changes) accounts for approximately 96.6% of the variance in y. The cyclical variable (unemployment) accounts for 2.1% of the variance observed, while the trend variable (program changes and the like) accounts for less than 1% of the variance.

Thus, when the effects of quarterly change and unemployment have been accounted for, there is only a slightly negative trend (6.6 FTE/QTR), which is not significantly different from zero, to relate to program changes. In short, it appears that the addition of new programs over the past twenty quarters has had no net effect on total University enrollments.

Unemployment, on the other hand, appears to significantly and positively impact total University enrollment. The data suggest that a 1% increase in Dade County unemployment will give rise to a 95 FTE increase in University enrollment the following quarter. Similarly, a decrease in the unemployment rate will cause a drop in enrollment. The correlation of these variables is clearly seen in figure 2 where deseasonalized enrollment and unemployment are plotted over time. The apparent increase in enrollment between Fall '73 to Fall '75 and the subsequent fall off of enrollment between Winter '76 and Spring '77 are substantively accounted for by a corresponding increase in Dade County unemployment from 4% to 12%, followed by a drop back to 6%, during the same periods.

The opening of the North Miami Campus also appears to have had a significant, positive impact on total University

enrollment. The model suggests an increase of 368 FTE students for the last five quarters, coincidental with the opening and operation of the North Miami Campus.

The real test of any model, of course, lies in its ability to predict future enrollments within a margin of error sufficiently low enough for planners to act upon the numbers generated with some degree of confidence. To make such predictions requires the application of several assumptions. First, use of a prediction model based on empirically derived relationships presumes these relationships will continue unchanged or, if changes occur, they can be systematically accounted for. Secondly, prediction of the dependent variable presumes that one can predict the independent variables. The accuracy and reliability of the dependent variable can be no better than the accuracy and reliability of the independent variables. Under normal circumstances, the further into the future one wishes to predict, the less accurate and reliable the estimate will be. Considering these independent variables in turn reveals that the trend factor has had very little, if any, impact on our enrollment. In predicting, it is assumed that this will continue to be the case. The factors which may affect the trend are not well understood but may be partially clarified when, at sometime in the future, we examine enrollment patterns by level and/or by school/college, program, etc. At least within the short term, there is no reason to expect a radical change in the trend relationship to the overall model.

The seasonal factors, the adjustment for regular quarterly changes, appears to remain fairly constant over time. Although slight shifts in these factors do occur, they can be systematically accounted for in a reasonably accurate and reliable manner.

The cyclical adjustment for unemployment, because we use previous quarter unemployment data, permits us to predict at least one quarter with reasonably high reliability and additional quarters to the extent the predictions of unemployment are reliable and accurate.

The impact of the North Miami Campus is more of an unknown. The use of a dummy variable for the North Miami Campus, while a reasonable technique for the short term, will not necessarily account for changes or growth that may occur as new programs, more course sections, etc., are added to the campus. However, a separate trend variable can replace the dummy variable in the future.

In order to formulate a projection of the University's likely annual enrollment, averaged by quarter, each of the independent variables was estimated in the manner described

above and applied to the regression formula. This resulted in estimates of enrollments for quarters 21, 22, 23. The average of these estimates taken with the actual quarter 20 enrollments (needed to make one full academic year) equals 6,340 FTE. As of February, 1979 the most likely actual annual enrollment averaged by quarter is approximately 6500 F.T.E. The projection model is in error by slightly less than 3%, on an annual basis.

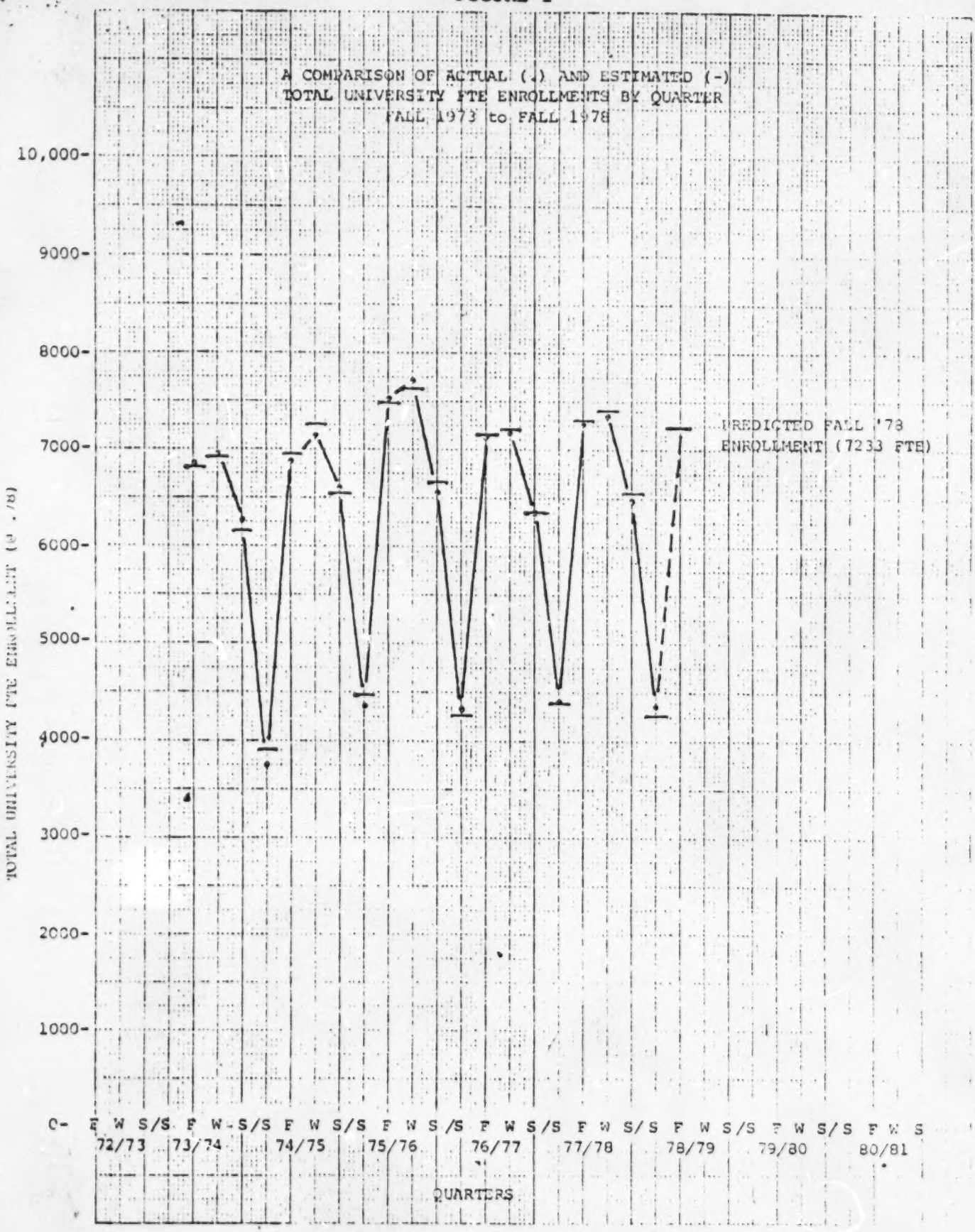
Conclusion

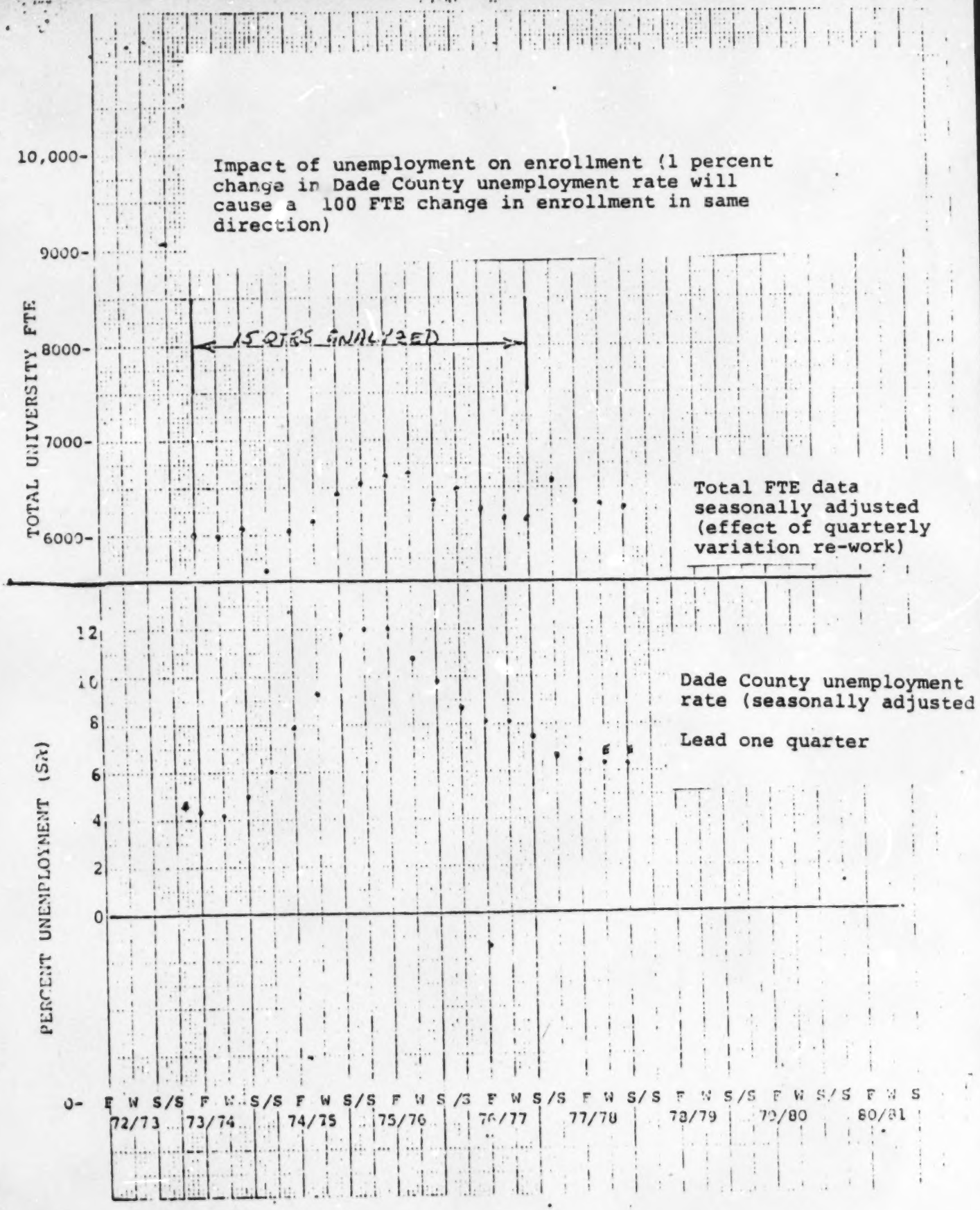
A regression model has been generated utilizing as independent variables a set of factors which can be estimated up to one year in advance of the year to be predicted. Results to date indicate that the procedures implemented yield predictions with residual error within a tolerable range for planning purposes. The explanatory nature of the regression coefficients have been exploited to add to our institution's understanding of the factors accouting for our apparent enrollment fluctuation. It is suggested that other institutional researchers may find the techniques herein discussed similarly useful.

REFERENCES

Salley, Charles, "Predicting Next Year's Resources -- Short Term Enrollment Forecasting for Accurate Budget Planning", paper presented to the Association for Institutional Research Annual Forum, May, 1978.

FIGURE I


A COMPARISON OF ACTUAL (.) AND ESTIMATED (-)
TOTAL UNIVERSITY FTE ENROLLMENTS BY QUARTER
FALL 1973 to FALL 1978
10,000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
TOTAL UNIVERSITY FTE ENROLLMENT (F '73 - F '78)
PREDICTED FALL '78
ENROLLMENT (7233 FTE)
F W S/S F W S/S F W S/S F W S/S F W S/S F W S/S F W S/S F W S/S F W S/S
72/73 73/74 74/75 75/76 76/77 77/78 78/79 79/80 80/81
QUARTERS


Impact of unemployment on enrollment (1 percent change in Dade County unemployment rate will cause a 100 FTE change in enrollment in same direction)
10,000-
9000-
8000-
7000-
6000-
TOTAL UNIVERSITY FTE
15 QTRS ANALYZED
Total FTE data seasonally adjusted (effect of quarterly variation re-work)
12
10
8
6
4
2
0
PERCENT UNEMPLOYMENT (SA)
Dade County unemployment rate (seasonally adjusted
Lead one quarter
F W S/S F W S/S F W S/S F W S/S F W S/S F W S/S F W S/S F W S/S F W S/S F W S
72/73
73/74
74/75
75/76
76/77
77/78
78/79
79/80
80/81

DATA SUMMARY

YEAR	TERM	Y OBS.	X_1	X_2	X_3	X_4	Y EST.	Y OBS. - Y EST.
1973	FALL	6847.9	1	1.140	4.3	0	6807.0	40.9
1974	WINTER	6958.8	2	1.161	4.2	0	6924.3	34.5
1974	SPRING	6263.7	3	1.029	4.9	0	6145.1	118.6
1974	SUMMER	3752.9	4	0.670	5.0	0	3865.8	(112.9)
1974	FALL	6879.8	5	1.140	6.0	0	6941.9	(62.1)
1975	WINTER	7143.6	6	1.161	7.9	0	7249.3	(105.7)
1975	SPRING	6614.5	7	1.029	9.3	0	6536.6	77.9
1975	SUMMER	4376.1	8	0.670	11.8	0	4485.3	(109.2)
1975	FALL	7532.4	9	1.140	12.0	0	7485.4	47.0
1976	WINTER	7718.6	10	1.161	12.0	0	7612.3	106.3
1976	SPRING	6556.4	11	1.029	10.8	0	6652.5	(96.1)
1976	SUMMER	4349.1	12	0.670	9.8	0	4268.7	80.4
1976	FALL	7141.8	13	1.140	8.7	0	7145.4	(3.6)
1977	WINTER	7191.6	14	1.161	8.1	0	7215.2	(23.6)
1977	SPRING	6377.0	15	1.029	8.1	0	6369.5	7.5
1977	SUMMER	4419.3	16	0.670	7.5	1	4391.8	27.5
1977	FALL	7259.3	17	1.140	6.7	1	7297.0	(37.7)
1978	WINTER	7369.9	18	1.161	6.5	1	7404.9	(35.0)
1978	SPRING	6494.1	19	1.029	6.4[e]	1	6549.6	(55.5)
1978	SUMMER	4352.0	20	0.670	6.3[e]	1	4251.3	100.7
1978	FALL	?	21	1.140	6.3[e]	1	7233.0	?

Y OBS = Actual Total University FTE enrollments @ .78

X_1 = Trend factor (quarter)
X_2 = Seasonal factor (quarterly variation adjustment)
X_3 = Cyclical factor - Dade County unemployment rate (seasonally adjusted) lead one quarter.
X_4 = Dummy variable for North Miami Campus

Y EST = Estimated Total University FTE enrollment
(Y EST = - 841.9 - 6.6 X_1 + 6357 X_2 + 95 X_3 + 368.2 X_4)

Y OBS - Y EST = residuals